

doi:10.12068/j.issn.1005-3026.2024.11.010

基于改进YOLOv8的铆接孔表面缺陷检测

郝博^{1,2}, 徐新岩¹, 赵玉欣¹, 闫俊伟¹

(1. 东北大学 航空动力装备振动及控制教育部重点实验室, 辽宁 沈阳 110819;

2. 东北大学秦皇岛分校 控制工程学院, 河北 秦皇岛 066004)

摘要: 飞机蒙皮、尾翼等零件上的铆接孔表面质量关乎飞机整体装配性能, 目前铆接孔缺陷检测大多为传统人工检测, 易出现漏检现象。因此, 提出一种具有创新性的改进YOLOv8的铆接孔表面缺陷检测方法。采用可变形卷积替换常规卷积, 解决特征提取中感受野形状固定的问题。嵌入SimAM注意力机制, 增强网络在背景和目标对比度较低状况下的辨识能力。使用WIoU边界框回归损失函数代替CIoU损失函数, 降低低质量图像对模型训练的影响, 提高模型的鲁棒性和泛化能力。为验证本文模型的性能, 以带铆接孔的6061铝合金板代替飞机蒙皮进行检测。实验结果表明, 本文模型在铆接孔测试集上mAP_{0.5}和准确率分别达到了0.918和0.920, 较原始YOLOv8模型分别提高了24.1%和25.3%。

关键词: YOLOv8; 铆接孔; 缺陷检测; 可变形卷积; 注意力机制

中图分类号: TP 391 文献标志码: A 文章编号: 1005-3026(2024)11-1595-09

Surface Defect Detection of Riveting Holes Based on Improved YOLOv8

HAO Bo^{1,2}, XU Xin-yan¹, ZHAO Yu-xin¹, YAN Jun-wei¹

(1. Key Laboratory of Vibration and Control of Aero-Propulsion System, Ministry of Education, Northeastern University, Shenyang 110819, China; 2. School of Control Engineering, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China. Corresponding author: XU Xin-yan, E-mail: 1035011108@qq.com)

Abstract: The surface quality of riveting holes on aircraft skin, tail and other components is crucial to the overall assembly performance of an aircraft. Currently, most riveting hole defect detection relies on the traditional manual methods, which are prone to missing detection. An improved detection method based on YOLOv8 for surface defect detection of riveting holes was proposed. The conventional convolution was replaced by deformable convolution to solve the problem of the fixed receptive field shape in feature extraction. The SimAM attention mechanism was embedded in order to enhance the recognition ability of the network under low contrast between the background and targets. The CIoU loss function was replaced by the WIoU bounding box regression loss function to reduce the impact of low-quality images during model training and improve the robustness and generalization ability of the model. To verify the performance of the improved model, 6061 aluminum alloy plates with riveting holes were used as a substitute for aircraft skin in the detection process. Experimental results demonstrated that the improved model achieved mAP_{0.5} and accuracy of 0.918 and 0.920 on the riveting hole test set, which represents an improvement of 24.1% and 25.3% compared to the original model.

Key words: YOLOv8; riveting hole; defect detection; deformable convolution; attention mechanism

飞机装配过程中机身蒙皮、尾翼等关键部位需要大量铆钉连接^[1]。铆接孔周围应力集中, 铆接

孔质量极大地影响铆接部位的疲劳寿命, 当铆接孔表面存在严重缺陷时会导致连接部位失效造

收稿日期: 2023-06-21

基金项目: 装备预先研究领域基金资助项目(61409230125)。

作者简介: 郝博(1963-), 男, 辽宁沈阳人, 东北大学教授, 博士生导师。

成飞机失事. 钻孔过程中存在因为设备性能不稳定、工人缺乏经验、劳动强度大等因素导致铆接孔表面出现凹陷、划痕等缺陷, 及时发现并进行修复是保障飞机装配质量的重要环节. 传统铆接孔表面缺陷检测依靠质检员肉眼观察, 检测效率低, 准确性难以得到保证, 因此提出一种高效率、高精度的铆接孔表面缺陷检测方法对飞机装配工作具有重要意义.

近年来, 深度学习技术在机器视觉目标检测领域发展迅猛, RCNN(regions with convolutional neural network)系列算法和 YOLO(you only look once)系列算法分别作为经典的双阶段和单阶段目标检测算法被国内外学者广泛应用于零件表面缺陷检测. 吕帅帅等^[2]改进 Mask-RCNN 算法, 提出基于铆钉位置反馈至裂纹定制化网络的机制, 并在 RPN(region proposal network)中增添 professional NMS(non-maximum suppression)算法对裂纹的 ROI(region of interesting)进行筛选, 该方法对飞行器的铆钉、裂纹识别准确率分别为 100% 和 87.5%. Liu 等^[3]设计一种基于改进 Faster RCNN 的航空发动机叶片表面缺陷检测系统, 采用 ROI Align 代替 ROI Pooling, 将特征金字塔网络与 ResNet-50 结合进行特征提取, 网络 mAP(mean average precision)达到了 79.0%. 潘睿志等^[4]提出一种多尺寸汽车轮辋焊缝检测与定位系统, 通过引入 CBAM(convolutional block attention module)混合域注意力机制、采用 EIoU 边框位置回归损失函数、增加 ASFF(adaptively spatial feature fusion)自适应特征融合网络对 YOLOv5 算法进行改进, 其准确率和 mAP_{0.5} 分别达到了 98.4% 和 99.2%. Aboah 等^[5]提出少镜头数据采样技术并利用 YOLOv8 模型实时检测多类头盔违规使用情况, 其 mAP 值为 0.586 1. 对比双阶段、单阶段网络在零件表面缺陷检测领域的应用, 单阶段网络不仅检测精度更高, 检测速度也有明显优势.

因铆接孔表面缺陷与背景纹理特征相似、同类缺陷尺寸大小不一等因素导致缺陷识别存在一定困难, 因此提出改进 YOLOv8 的铆接孔表面缺陷检测方法. 该方法以 YOLO 系列中最先进的 YOLOv8 模型为基础, 对网络结构和损失函数进行优化, 实现对缺陷目标的高准确率识别. 在主干网络中采用可变形卷积替换常规卷积, 以适应不同形态的缺陷. 在颈网络中引入 SimAM 注意力机制, 提高网络对背景信息的抗干扰能力. 使

用 WIoU 边界框回归损失函数代替 CIoU 损失函数, 减小低质量图像对模型训练的影响, 提高模型的鲁棒性和泛化能力.

1 铆接孔图像采集与预处理

本文使用维视 MV-EM510M 型 CCD 相机搭配 BT-R23C144 超紧凑型双远心镜头进行铆接孔表面图像采集, 相机与镜头设备参数如表 1, 表 2 所示. 远心镜头的成像原理为平行光成像, 畸变率极低, 通常小于 0.08%, 大大提高了拍摄精度. 选用 LED 环形光源对布满半径为 2mm 铆接孔的 6061 铝合金板打光, 该类型光源具有光束集中、均匀性好等优势. 图像采集设备见图 1a, 采集的铆接孔图像见图 1b.

表 1 CCD 相机参数

Table 1 Parameters of the CCD camera

型号	MV-EM510M
分辨率	2 456×2 058
外形尺寸/mm	50×50×48
最大帧率/(帧·s ⁻¹)	15
曝光方式	帧曝光
功耗/W	2.5

表 2 远心镜头参数

Table 2 Parameters of the telecentric lens

型号	BT-R23C144
放大倍率	0.061
视场尺寸/mm	144.3×108.2
光圈	8
远心度/(°)	<0.05
景深/mm	90
畸变/%	<0.08

使用 Adobe Photoshop 软件对采集的铆接孔图像进行裁剪, 去除铝合金板上的多余像素, 最终裁剪为 300 像素×300 像素的铆接孔图像, 用于制作数据集.

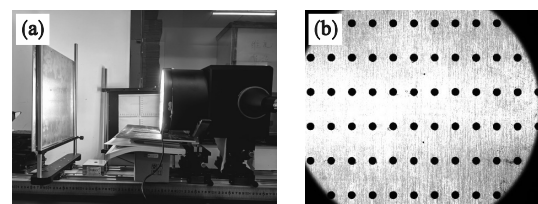


图 1 图像采集设备和铆接孔

Fig. 1 Image acquisition equipment and riveting holes

(a)—采集设备; (b)—铆接孔.

2 YOLOv8 网络及改进

2.1 YOLOv8 算法原理

YOLOv8 于 2023 年 1 月由开发 YOLOv5 的公司 Ultralytics 发布. YOLOv8 算法以 Darknet-53 为基础, 其实现特征提取的主干网络相比于 YOLOv5 算法, 使用梯度流更丰富的 C2f 模块代替 C3 模块, 使模型更轻量化. C2f 模块使用 chunk 函数将 CBS(Conv+BN+SiLU) 输出拆分成两个张量, 通过并行更多残差模块梯度流分支, 实现更高的检测精度.

YOLOv8 网络的预测端使用目前主流的解耦头结构代替耦合头结构, 分别进行分类与回归运算, 再整合预测结果. 分类任务提取类别信息时更关注目标内容, 而回归任务提取位置信息时更关注目标边缘, 因此, 解耦头结构使模型收敛速度更快, 降低模型训练的时间成本.

YOLOv8 采用 Anchor free 的方式, 摆脱了对

锚框的依赖, 解决了大部分锚框分布在背景区域上导致模型训练时正负样本分配严重不均衡的问题. 相较于 Anchor base, Anchor free 无须设置超参数来限定锚框的尺寸、宽高比与数量, 在一定程度上提高了目标检测模型的泛化能力.

在标签分配策略方面, YOLOv8 引用了任务对齐单阶段目标检测^[6](task-aligned one-stage object detection, TOOD) 的 TAL(task alignment learning) 动态标签匹配策略, 摒弃了 YOLOv5 的静态标签分配策略, 提高了分类和回归的准确率.

在损失函数方面, YOLOv8 预测端的损失包括类别预测损失和边界框回归损失两部分, 删减了置信度损失. 通过引入分布焦点损失(distribution focal loss, DFL) 函数^[7], 使网络更快聚焦到目标位置, 获取目标位置坐标.

2.2 改进 YOLOv8 算法原理

改进 YOLOv8 网络结构分为 4 部分: 输入端(Input)、主干网络(Backbone)、颈网络(Neck)、预测端(Prediction), 如图 2 所示.

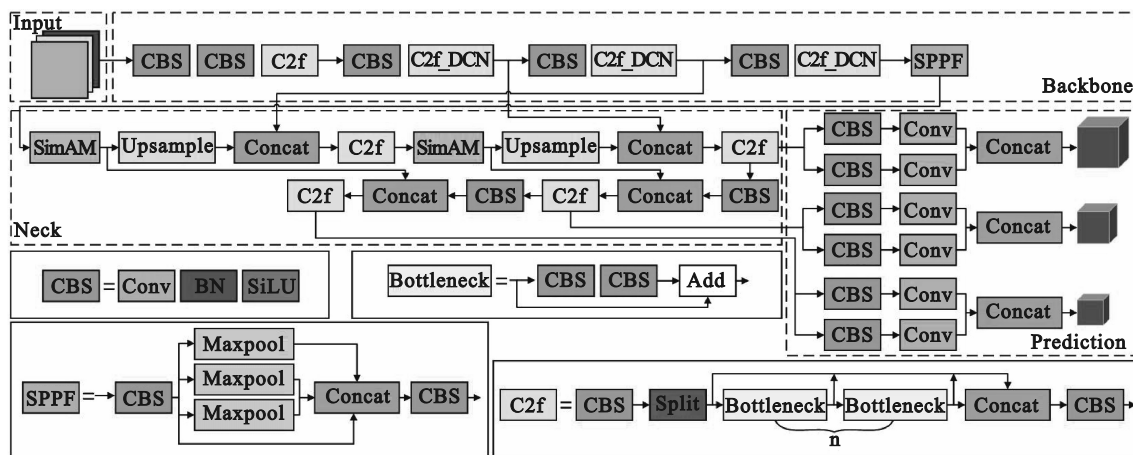


图 2 改进 YOLOv8 网络结构

Fig. 2 Network structure of the improved YOLOv8

2.2.1 可变形卷积替换常规卷积

常规卷积是卷积核在输入特征图上逐行逐列滑动, 将卷积矩阵和输入特征图的像素矩阵相乘求和, 进而得到输出特征图. 但常规卷积在固定位置对特征图进行采样, 感受野始终为矩形, 因此无法适应目标形变, 导致提取的目标关键信息缺失, 模型的泛化能力下降. 常规卷积如式(1)所示:

$$y(p_0) = \sum_{p_n \in R'} w(p_n) \cdot x(p_0 + p_n). \quad (1)$$

式中: x 为输入特征图; y 为输出特征图; p_0 为输出特征图中每个位置; p_n 为卷积核采样网格点 R' 中

所有采样位置; $w(p_n)$ 为采样位置的卷积核权重.

为进一步提升网络性能, 摆脱固定矩形感受野的局限性, 引入可变形卷积(deformable convolution)代替常规卷积. 可变形卷积通过引入偏移量来实现对输入特征图的局部采样, 从而适应不同形状和大小的目标, 计算流程如图 3 所示.

由图 3 可知, 可变形卷积结构是在常规卷积结构上增加了一个计算输入特征图偏移量的卷积分支, 得到每个采样点的偏移量. 因偏移量数值较小, 需利用双线性插值方法处理, 最终得到输出特征图. 可变形卷积如式(2)所示:

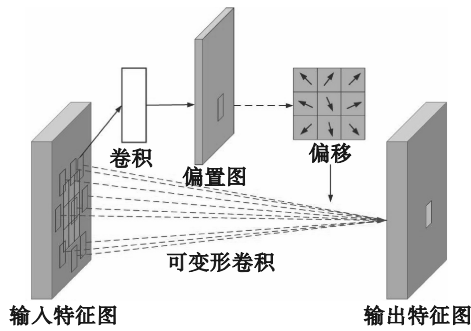


图3 可变形卷积计算流程

Fig. 3 Calculation flowchart of deformable convolution

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (2)$$

式中, Δp_n 为偏移量.

偏移后的采样点像素通过双线性插值计算, 解决了反向传播问题, 如式(3)所示:

$$x(p) = \sum_q G(q, p) \cdot x(q) \quad (3)$$

式中: $p = p_0 + p_n + \Delta p_n$ 为任意位置; q 为 p 周围的 4 个整数点; 二维双线性插值核函数 G 可拆分为两个一维核, 如式(4)、式(5)所示:

$$G(q, p) = g(q_x, p_x) \cdot g(q_y, p_y), \quad (4)$$

$$g(a, b) = \max\{0, 1 - |a - b|\}. \quad (5)$$

为防止可变形卷积神经网络^[8](deformable convolutional network, DCN)引入无用上下文, 干扰目标的特征提取, 本文增加幅度调节机制, 使用更多的可变形卷积结构, 对带偏移量的采样点赋予权重, 保证了目标信息的准确提取^[9], 如式(6)

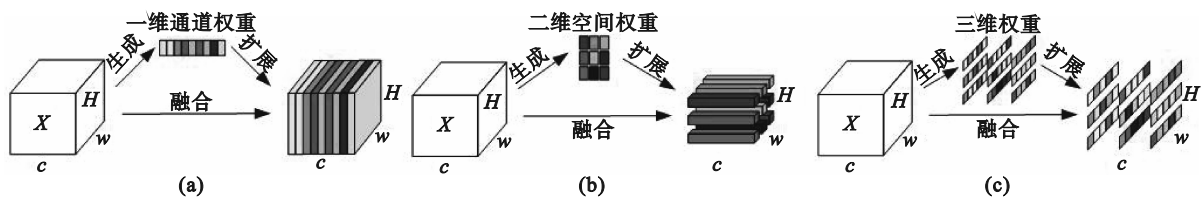


图4 注意力机制对比图

Fig. 4 Comparison of attention mechanisms

(a)一通道注意力机制; (b)一空间注意力机制; (c)一SimAM注意力机制.

受人脑注意力机制的启发, 含丰富信息的神经元会较其他神经元表现更为显著, 会对周围神经元产生抑制作用, 即空域抑制, 而这些具有空域效应的神经元应赋予更高的权重. SimAM作为一种 3D 无参注意力机制可通过定义能量函数直接评估每个独立神经元的重要性, 如式(7)所示:

$$e_i(w_i, b_i, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_i x_i + b_i))^2 + (1 - (w_i t + b_i))^2 + \lambda w_i^2 \quad (7)$$

式中: t 为目标神经元; x_i 为同一通路上其他神经

所示:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \cdot \Delta m_n \quad (6)$$

式中, Δm_n 为调制权重, 且 $0 \leq \Delta m_n \leq 1$.

由于铆接孔的同类型表面缺陷存在尺寸、面积大小不一的情况, 而常规卷积采用固定的矩形几何结构, 对目标检测有一定的局限性. 因此, 本文将特征提取网络中的第二至第四个 C2f 模块中的常规卷积网络替换为 DCN, 增强了模型对目标关键特征信息的捕捉能力.

2.2.2 SimAM 注意力机制

嵌入注意力机制能够提升网络的自适应注意力, 典型注意力机制包括 SENet (squeeze and excitation networks) 网络^[10]、CBAM (convolutional block attention module) 网络^[11]、BAM (bottleneck attention module) 网络^[12]等. SENet 网络包含通道注意力模块, 对输入特征层取全局平均池化, 再进行两次全连接, 获取所有通道信息, 并自主学习不同通道特征的重要程度. CBAM 包含通道注意力和空间注意力模块, 两模块串行组合, 同时关注每个通道的比重和每个特征点的比重, 考虑了全局上下文信息. BAM 将通道注意力和空间注意力模块并行组合, 并采用空洞卷积扩大空间注意力子模块的感受野, 有效利用了上下文信息.

然而 SimAM^[13]与上述注意力机制不同, 并非简单地将通道注意力和空间注意力模块串联或并联, 注意力机制对比如图 4 所示.

元; $w_i x_i + b_i$ 和 $w_i t + b_i$ 分别为 x_i, t 的线性变换; M 为每个通道能量函数的个数; λ 为超参数.

经计算 w_i, b_i 的解析解及除目标神经元外的其他神经元在通道维度上的均值和方差, 得到最小能量, 如式(8)所示:

$$e_i^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (8)$$

$$\text{式中: } \hat{\mu} = \frac{1}{M} \sum_{i=1}^M x_i; \hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{\mu})^2.$$

通过式(8)度量目标神经元和其他神经元之间的线性可分性, e_i^* 值越小, 表明神经元的重要程度越高. SimAM 最终优化为

$$\hat{X} = \text{sigmoid}\left(\frac{1}{E}\right) \odot X. \quad (9)$$

式中, E 将通道和空间维度的能量函数 e_i^* 分组, 使用 sigmoid 函数进行特征增强并限制 E 中较大的值.

本文提出的改进模型在颈网络中引入 SimAM 注意力机制, 相较于 SENet, CBAM 等, 其无需引入额外学习参数, 可以更全面考量特征的重要性, 增强铆接孔表面缺陷特征, 减弱背景纹理干扰特征, 提高模型的目标检测能力.

2.2.3 基于动态非单调聚焦机制的边界框损失

YOLOv8 模型训练损失包括类别预测损失 L_{cls} 和边界框回归损失 L_{box} , 类别预测依旧采用 BCE Loss 函数, 而边界框回归采用 CIoU Loss 函数和 DFL 函数相结合的形式.

CIoU Loss 考虑到了重叠面积、中心点距离、长宽比共 3 种重要几何参数, 如式(10)所示:

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{(\text{gt})})}{c^2} + \alpha v. \quad (10)$$

式中: α 为权重系数见式(11); v 为长宽比衡量函数, 见式(12); $b, b^{(\text{gt})}$ 分别表示预测框和真实框的中心点坐标; ρ 为两点间欧氏距离; c 为同时包含预测框和真实框的最小矩形对角线长度.

$$\alpha = \frac{v}{1 - \text{IoU} + v}, \quad (11)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{(\text{gt})}}{h^{(\text{gt})}} - \arctan \frac{w}{h} \right)^2. \quad (12)$$

DFL 以交叉熵的形式优化网络, 提升 y_i, y_{i+1} 接近目标 y 的概率, 从而让网络更快地聚焦到目标 y 的邻近区域, 其全局最小解能确保回归估计值无限接近目标坐标值, 如式(13)所示:

$$\text{DFL}(S_i, S_{i+1}) = -((y_{i+1} - y) \lg(S_i) + (y - y_i) \lg(S_{i+1})). \quad (13)$$

式中: S_i, S_{i+1} 分别为 y_i, y_{i+1} 的广义分布概率值; $y_i \leq y \leq y_{i+1}$.

但若目标检测训练集中含低质量样本, 强化边界框回归能力会降低模型的检测性能. 因此, 本文采用基于动态非单调聚焦机制 (Wise-IoU, WIoU) 的边界框损失 WIoU Loss^[14] 代替 CIoU Loss. 动态非单调聚焦机制通过使用离群度及梯度增益分配策略, 可以聚焦于普通质量的锚框, 并有效削弱低质量样本对锚框回归的影响, 提高

模型的鲁棒性和泛化能力. WIoU 共 3 个版本: 具有两层注意力机制的 WIoU v1; 具有单调聚焦机制的 WIoU v2; 具有动态非单调聚焦机制的 WIoU v3, 本文采用 WIoU v3, 如式(14)所示:

$$L_{\text{WIoU}_v3} = r \times \exp\left(\frac{(x - x_{\text{gt}})^2 + (y - y_{\text{gt}})^2}{(W_{\text{g}}^2 + H_{\text{g}}^2)^*}\right) \times (1 - \text{IoU}). \quad (14)$$

式中: r 为梯度增益见式(15); β 为离群度; α, δ 为超参数; $(x, y), (x_{\text{gt}}, y_{\text{gt}})$ 分别为预测框和真实框的中心坐标; $W_{\text{g}}, H_{\text{g}}$ 分别为同时包含预测框和真实框的最小矩形的长、宽. 本文设置 $\alpha = 1.9, \delta = 3$.

$$r = \frac{\beta}{\delta \alpha^{\beta - \delta}}. \quad (15)$$

改进后总损失函数按权重比求和, 如式(16)所示:

$$\text{Loss}_{\text{Total}} = w_1 L_{\text{cls}} + w_2 L_{\text{WIoU}} + w_3 L_{\text{D}}. \quad (16)$$

式中, L_{D} 为 DFL(S_i, S_{i+1}), 本文设置权重 $w_1 = 0.5, w_2 = 7.5, w_3 = 1.5$.

3 实验和结果分析

3.1 实验设计与评价指标

实验操作系统为 Windows11, CPU 型号为 Intel Core i9-13900HX 2.20GHz, 运行内存 16 GB, GPU 型号为 NVIDIA GeForce RTX 4060, 总显存 15.9 GB, GPU 加速库为 CUDA11.8 和 CUDNN8.7.0, 深度学习框架为 Pytorch2.0.0, 编程语言为 Python3.8, 开发软件为 PyCharm Community Edition.

在目标检测领域, 卷积神经网络能够取得优异的效果依赖于有监督的训练, 其需要足够多的训练样本. 少样本易出现模型过拟合、检测性能差等情况, 本文采集铆接孔缺陷样本 320 个, 采用旋转、平移、翻转、调整明亮度、添加噪声(椒盐、高斯噪声)等多种数据扩增方式提升样本规模, 得到 1 280 张铆接孔表面缺陷图像. 对扩增后的数据集, 按照 6:2:2 的比例划分为训练集、验证集和测试集, 使用人工标注工具 LabelImg 对铆接孔缺陷进行标注, 图 5 为可视化数据集 4 类缺陷样本标签的数量, 其中包括 384 个裂纹 (crack)、576 个凹陷 (sunken)、256 个毛刺 (burr)、256 个划痕 (scratch). 图 6 为数据集标签分布及尺寸情况, 标签框的中心点坐标、宽度、高度均经过归一化处理, 缺陷目标主要集中在图像中央位置, 小目标较多.

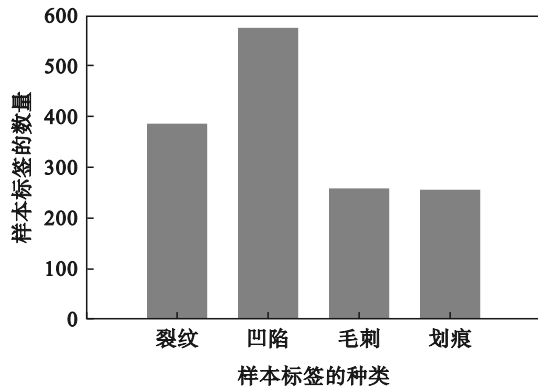


图5 数据集样本的种类和数量

Fig. 5 Types and quantities of the data set samples

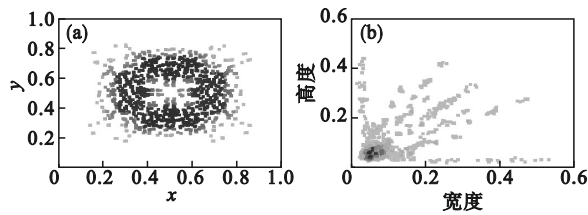


图6 数据集目标分布及尺寸情况

Fig. 6 Target distribution and size of the data set
(a)—数据集目标分布; (b)—尺寸.

选用精确率 P (precision)、召回率 R (recall)、准确率 A_{cc} (accuracy)、平均精度 AP (average precision)、平均精度均值 mAP、F1 值 ($F1$ -score) 等指标来评价铆接孔缺陷检测模型的性能. 上述评价指标可通过混淆矩阵计算, 计算如式(17)~式(21)所示:

$$P = TP / (TP + FP), \quad (17)$$

$$R = TP / (TP + FN), \quad (18)$$

$$A_{cc} = (TP + TN) / (TP + TN + FP + FN), \quad (19)$$

$$AP = \int_0^1 P(R) dR, \quad (20)$$

$$F1 = 2P \cdot R / (P + R). \quad (21)$$

式中: TP 为真阳性; FN 为假阴性; FP 为假阳性; TN 为真阴性.

3.2 模型训练

本文方法选择的训练模型为 YOLOv8s, 训练轮次为 1 300, 批样本数量为 16, 初始学习率和最终学习率均为 0.01, 动量为 0.937, 权值衰减为 0.000 5.

基于上述实验配置, 对比本文模型、YOLOv8 和 YOLOv5 在验证集上的训练结果. 训练过程中边界框回归损失和类别预测损失如图 7 所示.

边界框回归损失用于衡量模型预测的目标边界框与真实框间的差异, 类别预测损失用于衡量模型能否准确识别目标种类, 损失值越小, 模

型性能越好. 由图 7 可知, 随着训练轮次的增加, 3 种模型的损失值均逐渐下降并趋于稳定. 本文模型相比于 YOLOv8, YOLOv5 的损失值最小, 收敛速度最快.

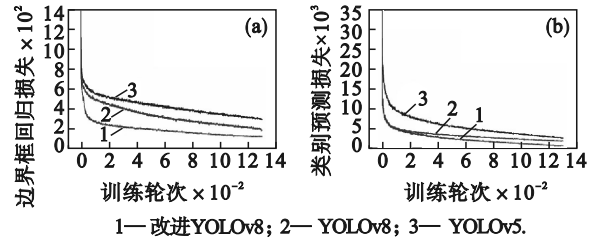


图7 不同模型损失值对比

Fig. 7 Loss value comparison of different models

(a)—边界框回归损失值; (b)—类别预测损失值.

图 8 为 3 种模型在验证集上 mAP 值对比, mAP_0.5 表示 IoU 阈值取 0.5 时的平均精度均值, mAP_0.5:0.95 表示在 IoU 阈值为 0.5 到 0.95、步长为 0.05 上的平均 mAP. 由图 8a 可知 YOLOv5, YOLOv8 对铆接孔表面缺陷检测的 mAP_0.5 值分别为 0.623 和 0.686, 本文模型的 mAP_0.5 值达到了 0.934. 由图 8b 可知 YOLOv5, YOLOv8 模型的 mAP_0.5:0.95 值分别为 0.315 和 0.381, 本文模型的 mAP_0.5:0.95 值达到了 0.531. 表明本文模型对铆接孔表面缺陷具有较好的检测效果.

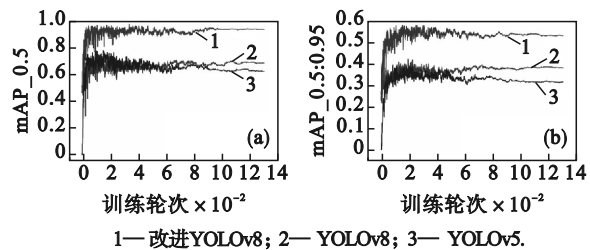


图8 不同模型 mAP 值对比

Fig. 8 mAP value comparison of different models

(a)—mAP_0.5; (b)—mAP_0.5:0.95.

3.3 模型评估与消融实验

为了更直观地评估本文模型的检测效果, 基于测试集进行分析, 对比改进前后 YOLOv8 模型的混淆矩阵, 如图 9 所示. 其中, 凹陷样本 96 个, 裂纹、毛刺及划痕样本均为 64 个.

混淆矩阵背景类的行数值表示漏检数量, 列数值表示虚检数量. 在铆接孔表面缺陷检测中, 铆接孔图像的背景与凹陷、划痕缺陷对比度低, 因此易出现漏检情况. 由图 9a 计算可知, YOLOv8 模型在测试集上的准确率 A_{cc} 为 66.7%, 漏检率为 29.5%, 虚检率为 5.0%, YOLOv8 模型准确率且

漏检严重.由图 9b 可知,本文模型的准确率 A_{cc} 为 92.0%,漏检率为 6.3%,虚检率为 6.2%.本文模型虽在虚检率方面略高于 YOLOv8 模型 1.2%,但其准确率提高 25.3%,漏检率降低 23.2%.

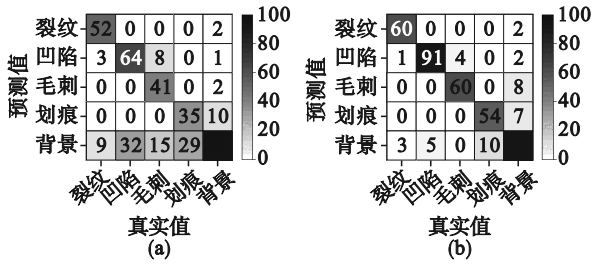


图 9 两种模型混淆矩阵图

Fig. 9 Confusion matrix of the two models

- (a)—YOLOv8 模型混淆矩阵;
- (b)—改进 YOLOv8 模型混淆矩阵.

通过混淆矩阵计算及精确率-召回率曲线 (PR 曲线) 对比分析,得出 4 类缺陷的精确率 P 、召回率 R 、 $F1$ 值、平均精度 AP,如表 3 所示.由表 3 可以确定本文模型对缺陷目标检测精度优于

YOLOv8 模型.

表 3 两种模型检测精度参数

Table 3 Test accuracy parameters of the two models

模型	缺陷	P	R	$F1$ 值	AP_0.5
本文模型	裂纹	0.919	0.938	0.928	0.912
	凹陷	0.900	0.948	0.923	0.963
	毛刺	0.856	0.922	0.888	0.929
	划痕	0.890	0.759	0.819	0.869
YOLOv8 模型	裂纹	0.791	0.812	0.801	0.723
	凹陷	0.736	0.667	0.700	0.719
	毛刺	0.870	0.629	0.730	0.812
	划痕	0.575	0.487	0.527	0.453

为了进一步验证本文模型对铆接孔表面缺陷检测的有效性,定量评估 WIoU 损失函数、SimAM 注意力机制、可变形卷积 DCN 对 YOLOv8 模型的优化作用,设计了消融实验,结果如表 4、图 10 所示.

表 4 消融实验结果

Table 4 Results of ablation experiments

组别	模型	P	R	mAP_0.5	mAP_0.5:0.95	帧率/(帧·s ⁻¹)
1	YOLOv8	0.743	0.649	0.677	0.367	243.90
2	YOLOv8+WIoU	0.791	0.650	0.754	0.458	204.08
3	YOLOv8+WIoU+SimAM	0.790	0.681	0.806	0.476	178.57
4	YOLOv8+WIoU+SimAM+DCN	0.891	0.892	0.918	0.524	147.06

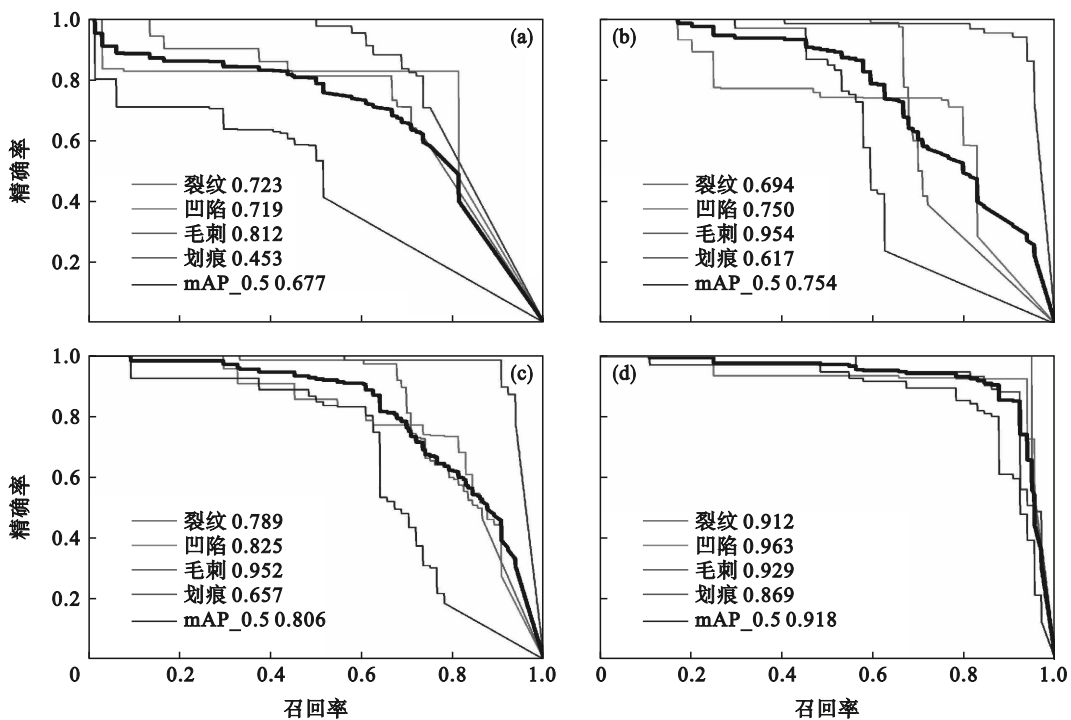


图 10 PR 曲线

Fig. 10 PR curves

- (a)—YOLOv8; (b)—YOLOv8+WIoU; (c)—YOLOv8+WIoU+SimAM; (d)—YOLOv8+WIoU+SimAM+DCN.

由表 4 可知,原始 YOLOv8 模型在测试集上的精确率、召回率、mAP_0.5、mAP_0.5:0.95 分别为 0.743, 0.649, 0.677, 0.367, 对比 1~4 组数据可以看出,每部分改进对网络性能的提升都作出了贡献.本文模型相较于 YOLOv8 模型在精确率、召回率、mAP_0.5、mAP_0.5:0.95 上分别提高了 14.8%, 24.3%, 24.1%, 15.7%, 但其每秒帧率 FPS (frames per second) 下降为 147.06 帧/s.改进后的模型虽检测速度有所降低,但检测精度大幅提高.

模型的 PR 曲线反映精确率 P 和召回率 R 的相互关系,曲线围成的面积即为 AP 值.由图 10a、图 10b 对比可知,当使用 WIoU 损失函数代替 CIoU 损失函数,模型对毛刺、划痕缺陷检测效果明显提高,AP 值分别提高了 14.2%, 16.4%.对图 10b、图 10c 分析可得,模型融入 SimAM 注意力机制后,削弱了背景纹理特征的影响,裂纹缺陷的 AP 值提高了 9.5%.由图 10c、图 10d 可以看出,可变形卷积替换常规卷积后,网络能自适应提取缺陷关键信息,进一步优化了模型的检测效果.

采用 Grad-CAM^[15]方法,可以清楚地反映模型改进前后对缺陷目标的关注程度,如图 11 所

示.相较于 CAM, Grad-CAM 不需要修改网络结构或重新训练就可以对卷积神经网络进行可视化,如式(22),式(23)所示:

$$L_{\text{Grad-CAM}}^c = \text{ReLU}\left(\sum_k \alpha_k^c A^k\right), \quad (22)$$

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}. \quad (23)$$

式中: A^k 为特征层 A 中通道 k 的数据; α_k^c 为 A^k 权重; y^c 为网络针对类别 c 预测的分数; A_{ij}^k 为特征层 A 在通道 k 中坐标 (i, j) 位置处的数据; Z 为特征层宽度与高度的乘积.

由图 11 可知,本文模型由于极少关注背景纹理特征,减少了背景对检测结果的干扰,同时提高对铆接孔表面缺陷特征的关注度,提高检测准确性,改善原始 YOLOv8 模型对缺陷定位不够准确的问题,说明本文提出的改进方案可行性较高.

图 12 为本文模型和 YOLOv8 模型的实际检测效果图,可知经过改进后的模型识别缺陷的置信度相比原始模型有较大提高.

由缺陷目标检测框定量分析可知,铆接孔数据集包含尺寸小于 30 像素×30 像素的缺陷样本,即微小缺陷样本.经上述实验验证,本文模型具有更强的微小缺陷检测能力.

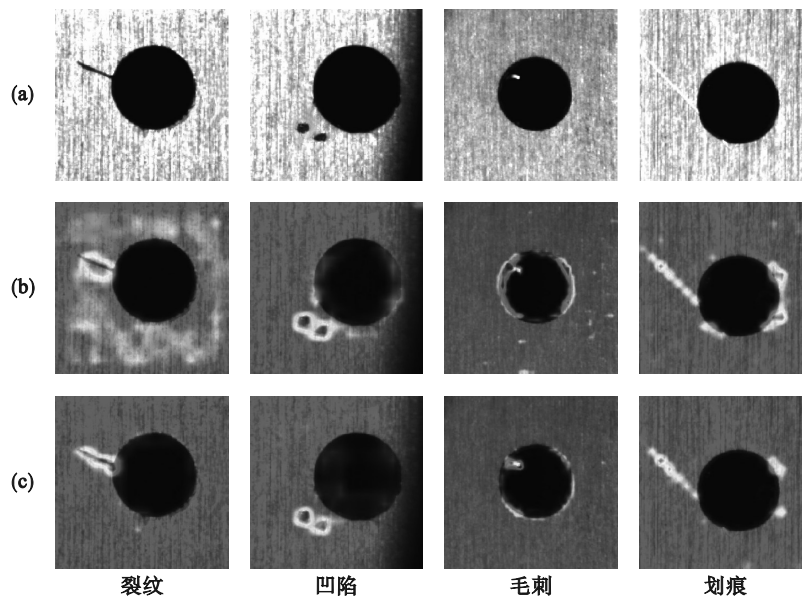
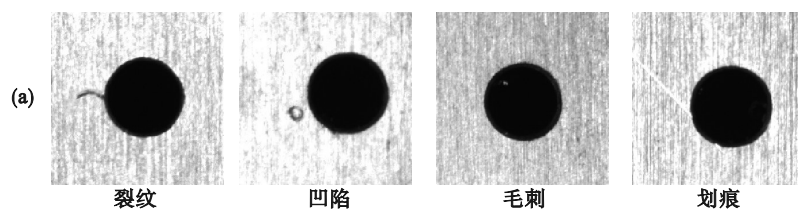


图 11 改进后模型与原始模型的热力图对比

Fig. 11 Comparison of heat maps between the improved model and the original model
(a)—缺陷原图;(b)—YOLOv8模型可视化热力图;(c)—改进 YOLOv8模型可视化热力图.



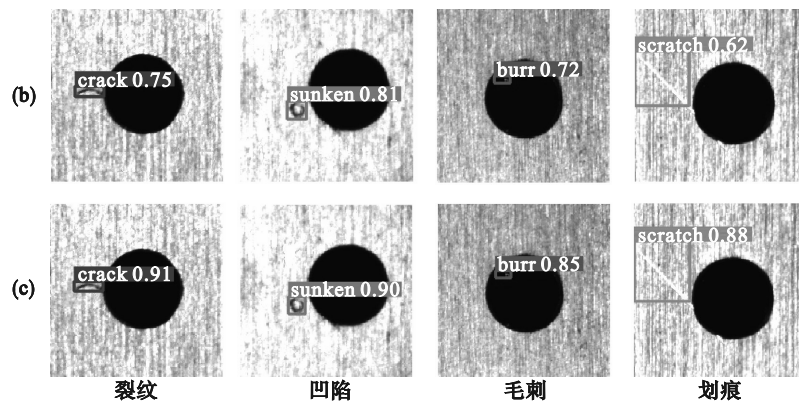


图 12 改进模型和原始模型的实际检测效果

Fig. 12 Actual detection effects of the improved model and the original model
(a)—缺陷原图; (b)—YOLOv8 模型检测结果; (c)—改进 YOLOv8 模型检测结果.

4 结 论

1) 以 YOLOv8 模型为基础,采用可变形卷积替换常规卷积,融入 SimAM 注意力机制,使用 WIoU 边界框回归损失函数代替 CIoU 损失函数,提升了模型对缺陷目标的检测能力,在验证集上 mAP_{0.5} 高达 0.934,相比于 YOLOv5 和原始 YOLOv8 模型分别提升了 31.1% 和 24.8%.

2) 基于测试集的消融实验结果表明,本文模型的帧率和 mAP_{0.5} 分别为 147.06 帧/s 和 0.918,虽检测速度有所降低,但检测精度大幅提高.由消融实验 PR 曲线对比可知,使用 WIoU 损失函数代替 CIoU 损失函数,模型对毛刺缺陷的 AP 值提高了 14.2%,对划痕缺陷的 AP 值提高了 16.4%.

3) 由测试集混淆矩阵计算可知,本文模型在虚检率方面略高于 YOLOv8 模型 1.2%,但其准确率提高 25.3%,漏检率降低 23.2%.由实际检测效果图看出,本文模型识别缺陷的置信度相比 YOLOv8 模型有较大提高.

参考文献:

- [1] Wang J, Zhu C R, Yang Y P, et al. Effect of riveting displacement on the mechanical behavior of CFRP bolted joints with elliptical-head non-lug self-locking rivet nut[J]. *International Journal of Advanced Manufacturing Technology*, 2023, 125(5/6): 2161–2182.
- [2] 吕帅帅,杨宇,王彬文,等.基于改进 Mask-RCNN 的飞行器结构裂纹自动检测方法[J].*振动、测试与诊断*, 2021, 41(3): 487–494, 620.
(Lyu Shuai-shuai, Yang Yu, Wang Bin-wen, et al. An automatic crack detection method for structure test based on improved Mask-RCNN [J]. *Journal of Vibration, Measurement & Diagnosis*, 2021, 41(3): 487–494, 620.)
- [3] Liu Y X, Wu D B, Liang J W, et al. Aeroengine blade surface defect detection system based on improved faster RCNN [J]. *International Journal of Intelligent Systems*, 2023: 1992415.
- [4] 潘睿志,林涛,李超,等.基于深度学习的多尺寸汽车轮辋焊缝检测与定位系统研究[J].*光学精密工程*, 2023, 31(8): 1174–1187.
(Pan Rui-zhi, Lin Tao, Li Chao, et al. Research on multi size automobile rim weld detection and positioning system based on depth learning [J]. *Optics and Precision Engineering*, 2023, 31(8): 1174–1187.)
- [5] Aboah A, Wang B, Bagci U, et al. Real-time multi-class helmet violation detection using few-shot data sampling technique and YOLOv8 [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Vancouver, 2023: 5350–5358.
- [6] Feng C J, Zhong Y J, Gao Y, et al. TOOD: task-aligned one-stage object detection [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, 2021: 3490–3499.
- [7] Li X, Wang W H, Wu L J, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection [J]. *ArXiv*, 2022: 2006.04388.
- [8] Dai J F, Qi H Z, Xiong Y W, et al. Deformable convolutional networks [C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, 2017: 764–773.
- [9] Zhu X Z, Hu H, Lin S, et al. Deformable ConvNets V2: more deformable, better results [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, 2019: 9300–9308.
- [10] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks [J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011–2023.
- [11] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision. Cham: Springer, 2018: 3–19.
- [12] Park J, Woo S, Lee J Y, et al. BAM: bottleneck attention module [J]. *ArXiv*, 2022: 1807.06514.
- [13] Yang L X, Zhang R Y, Li L D, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks [C]//International Conference on Machine Learning (ICML). San Diego, 2021: 1–12.
- [14] Tong Z J, Chen Y H, Xu Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [J]. *ArXiv*, 2023: 2301.10051.
- [15] Selvaraju R R, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization [J]. *International Journal of Computer Vision*, 2020, 128(2): 336–359.